# A Gaussian Mixture Method for Specific Differential Phase Retrieval at X-band Frequency

Guang Wen[1], Neil I. Fox[1,*], and Patrick S. Market[1]

[1]School of Natural Resources, University of Missouri, 332 ABNR Building, Columbia, Missouri, USA, 65201

**Correspondence:** Neil I. Fox (foxn@missouri.edu)

We would like to express our sincere thanks to the Editor for the efficient review management, and to the anonymous reviewers for their valuable comments and suggestions. We have addressed all the reviewers' comments point by point in the revision.

## 1 Review #1

**1.1.** ***Readers would benefit from a more tutorial style as the topic is highly specialized. This regards both the Kdp estimation in general as well as the Gaussian mixture statistical modeling.***

[response] This is really a good comment, since the paper describes a new $K_{dp}$ estimation method. It is necessary to give some details about this topic to readers in a variety of backgrounds. Therefore, we have provided an additional appendix related to the regression-based estimation of $K_{dp}$.

[changes] p.4, ln.10–12: This method has been widely used in the existing radar system (Cifelli et al., 2018; Chandrasekar et al., 2018; Chen et al., 2017b, a). The details of the regression-based estimation of $K_{dp}$ are given in Bringi and Chandrasekar (2001) and Appendix A.

[changes] p.19, ln.6–p.20, ln.6: **Appendix A: Regression-based estimation of $K_{dp}$**

Let the total differential phase $\psi_{dp}$ be $y$, and the range gate $r$ be $x$. The $\psi_{dp}$ profile over small range segments can be approximated by a first-order polynomial, i.e,i

$$y = \beta_0 + \beta_1 x + \epsilon, \tag{1}$$

where $\beta_0$ and $\beta_1$ are the coefficients in the linear approximation, and $\epsilon$ is an error function. It can be assumed that $\epsilon$ is independent and individual distributed with zero mean and variance of $\sigma_\epsilon^2 = \sigma^2$.

In the linear regression, it is easy to find that

$$\beta_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}. \tag{2}$$

where $\bar{x}$ and $\bar{y}$ are the means of $x$ and $y$ in the segment, respectively. Since

$$\sum_i (x_i - \bar{x})(y_i - \bar{y}) = \sum_i (x_i - \bar{x}) y_i - \sum_i (x_i - \bar{x}) \bar{y} \tag{3}$$

and

$$\sum_i (x_i - \bar{x})\bar{y} = \bar{y}\left(\sum_i x_i - N\bar{x}\right) = \bar{y}(N\bar{x} - N\bar{x}) = 0, \tag{4}$$

we have

$$\beta_1 = \frac{\sum_i (x_i - \bar{x}) y_i}{\sum_i (x_i - \bar{x})^2}, \tag{5}$$

where $N$ is the number of the gates in the segment.

It is noted that the range gate $r$ is equally spaced with an interval of $\Delta r$, $\psi_{dp}$ is the two-way propagation phase shift, and $K_{dp}$ is the one-way specific differential phase. The $K_{dp}$ is then estimated by

$$K_{dp} = \frac{\sum_{i=1}^{n} \psi_{dp}(r_i)\left[i - \frac{(n+1)}{2}\Delta r\right]}{\frac{1}{6} n(n-1)(n+1)\Delta r^2}. \tag{6}$$

At S-band, the backscattering differential phase shift $\delta_{co}$ is often negligible, and thus $\psi_{dp}$ and $\phi_{dp}$ are interchangeable, leading to Eq. (2).

By taking the variance on both sides of Eq. (5) and noting $\epsilon$ is the only variable, we have

$$\sigma^2(\beta_1) = \sigma^2\left(\frac{\sum_i (x_i - \bar{x})(\beta_0 + \beta_1 x_i + \epsilon)}{\sum_i (x_i - \bar{x})^2}\right) \tag{7}$$

$$= \frac{\sum_i (x_i - \bar{x})^2 \sigma_\epsilon^2}{\left[\sum_i (x_i - \bar{x})^2\right]^2} \tag{8}$$

$$= \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2} \tag{9}$$

Similar to Eq. (6), we have

$$\sigma^2(K_{dp}) = \frac{\sigma^2(\psi_{dp})}{\frac{1}{3}\Delta r^2 [n(n-1)(n+1)]}. \quad \# \tag{10}$$

[response] The Gaussian mixture model is widely used in signal processing, but may be new in atmospheric science. To interpret this model, we may think the mixture model is that the data looks multimodal, for example, a raindrop size distribution (DSD) with multiple peaks. Trying to fit a multimodal DSD with a unimodal model will lead to poor fitting. An obvious way to model a multimodal DSD would be to assume that it is generated by multiple unimodal DSD. In signal processing, a commonly used distribution is the Gaussian distribution. Therefore, modeling multimodal data as a mixture of many unimodel Gaussian distributions makes intuitive sense. We have added more words about Gaussian mixture model.

[changes] p.4, ln.25: Intuitively, it is used to model the multimodal data, with each Gaussian component corresponding to a subpopulation of the data.

**1.2.** ***I think, that it is important to point out that Kdp is calculated from a filtered (estimated) differential phase and not directly from its moment-based measurements.To distinguish the three, one may use psi, fi, k symbols.***

[response] It is true that we need to derive $\phi_{dp}$ from the raw data $\psi_{dp}$ before estimating $K_{dp}$, since the X-band radar is affected by the backscattering differential phase $\delta_{co}$. In fact, the Gaussian mixture method analyzes the raw $\psi_{dp}$ to calculate the mean $\psi_{dp}$ profile, and then remove $\delta_{co}$ to obtain the $\phi_{dp}$. In the revision, we have corrected the notation problems by denoting the data as the $\psi_{dp}$ before $\delta_{co}$ elimination and as the $\phi_{dp}$ after it.

[changes] We have made a number of changes:

p.7, ln.20–30: From the chart of LR in Fig. 1.a, we can see that after the radar measurements are collected, the $\psi_{dp}$ is unfolded, and then the clutter is removed. After these corrections, an iterative filtering method is applied to the $\psi_{dp}$ profile. An adaptive method is finally used to estimate the $K_{dp}$ profile according to the values of $Z_H$. The Gaussian mixture model, on the other hand, processes $\psi_{dp}$ differently. First of all, the clutter is masked out according to the thresholds of $Z_H$ and the variation of $\psi_{dp}$. Secondly, the range $r$ and $\psi_{dp}$ are fitted into a Gaussian mixture to yield the joint PDF, while the mean $\psi_{dp}$ and the $\psi_{dp}$ variance are obtained by taking the first raw and second central moments of the conditional PDF of $\psi_{dp}$ given $r$. Thirdly, some specific clusters in the Gaussian mixture PDF are adjusted to solve the problems of ambiguous $\psi_{dp}$ and backscattering differential phase shift $\delta_{co}$ in order to derive the PDF of $\phi_{dp}$. Fourthly, a raw $K_{dp}$ profile is calculated from the first derivative of the expected values of $\phi_{dp}$, and the associated variances are obtained via a Taylor series expansion. Finally, the raw $K_{dp}$ profile is smoothed, and consequently, the variances are reduced. In addition, new $\phi_{dp}$ with lower variances can be re-constructed from the $K_{dp}$ estimates.

**1.3.** ***The authors should emphasize that the main advantage of their proposed method is in providing the estimation variance for the Kdp and not is providing better estimates of Kdp. This is evident in the long-term evaluation using rain gauge data.***

[response] This is absolutely right that the Gaussian mixture method has the advantage that it provides the variance of $K_{dp}$ together with the mean $K_{dp}$. Since the $K_{dp}$ variance is nonconstant, it leads to the variability in the $K_{dp}$ error characteristics. Furthermore, the method yields the statistical uncertainty of $K_{dp}$, which is often missed in the existing methods. We can then use the uncertainty of $K_{dp}$ to calculate the uncertainty of $Z_H$ and $Z_{DR}$ via the attenuation correction, and the uncertainty of $R$ via the $R$–$K_{dp}$ relation. These uncertainties are useful for studying the streamflow trends in the hydrological model. It is true that our rain rate estimates are not optimized for the MZZU radar, since we did not derive the $R$–$K_{dp}$ relation in the paper. For the rain rate estimation, one can refer to some advanced studies, such as the IFloodS (Chen et al., 2017a) and MC3E (Giangrande et al., 2014) campaigns.

[changes] We have made a number of changes:

p.2, ln. 34–p.3, ln.3: It is found that $\sigma^2(K_{dp})$ is closely related to the square of the first derivative of $K_{dp}$ and $\sigma^2(\Phi_{dp})$, while large $\sigma^2(K_{dp})$ is associated with high variation of $K_{dp}$ estimates. When compared to the existing methods, our method considers the joint probability density function of the data as the non-linear Gaussian mixture, leading to better performance for the multimodal data. The $K_{dp}$ variance can be used to calculate the variances of $Z_H$, $Z_{DR}$ and rain rate, and to study the streamflow trends in the hydrological model.

p.17, ln. 20–25: It is clear that the rain rates based on the GMM $K_{dp}$ have a moderate consistency with the rain gauge data. To improve the results, some advanced rain rate algorithms can be considered, such as the rain-ice separation technique in the IFloodS campaign (Chen et al., 2017a) and the radar-gauge comparison method in the MC3E campaign (Giangrande et al., 2014). Nevertheless, the GMM has the advantage over the existing methods, since it can yield the variance of $K_{dp}$. Furthermore, the variance of $R$ can also be obtained by the mean $K_{dp}$ and the $K_{dp}$ variance via the $R$–$K_{dp}$ relation, leading to the variability in the error characteristics of $R$. Thus, the variances can be used to study the streamflow trends in the hydrological model.

**1.4. *I suggest that the authors improve the quality of the figures: some lettering is not legible, the inter-panel space could be reduced, etc.***

[response] We have improved the figures according to the suggestion.

[changes] Figures 1, 2, 4, 7, 8 and 9.

## References


Bringi, V. and Chandrasekar, V.: Polarimetric Doppler weather radar: Principles and applications, Cambridge Univ Pr, 2001.

Chandrasekar, V., Chen, H., and Philips, B.: Principles of High-Resolution Radar Network for Hazard Mitigation and Disaster Management in an Urban Environment, Journal of the Meteorological Society of Japan. Ser. II, advpub, 2018.

Chen, H., Chandrasekar, V., and Bechini, R.: An Improved Dual-Polarization Radar Rainfall Algorithm (DROPS2.0): Application in NASA IFloodS Field Campaign, Journal of Hydrometeorology, 18, 917–937, 2017a.

Chen, H., Lim, S., Chandrasekar, V., and Jang, B.-J.: Urban Hydrological Applications of Dual-Polarization X-Band Radar: Case Study in Korea, Journal of Hydrologic Engineering, 22, E5016 001, 2017b.

Cifelli, R., Chandrasekar, V., Chen, H., and Johnson, L. E.: High resolution radar quantitative precipitation estimation in the San Francisco Bay area: Rainfall monitoring for the urban environment, Journal of the Meteorological Society of Japan. Ser. II, 96, 141–155, 2018.

Giangrande, S. E., Collis, S., Theisen, A. K., and Tokay, A.: Precipitation Estimation from the ARM Distributed Radar Network during the MC3E Campaign, Journal of Applied Meteorology and Climatology, 53, 2130–2147, 2014.